

Ses Komut Tanıma ile Gezgin Araç Kontrolü

Rifat Edizkan¹, Burak Tiryaki¹, Tunç Büyükcan¹, İsmail Uzun²

¹ Eskişehir Osmangazi Üniversitesi, Elektrik Elektronik Mühendisliği Bölümü, 26480, Eskişehir

² EKOM İletişim Teknolojileri, TÜBİTAK-MAM Teknoloji Serbest Bölgesi, 41470, Gebze, Kocaeli

redizkan@ogu.edu.tr, btiryaki@ogu.edu.tr, tbuyukcan@ogu.edu.tr, iuzu@ekom.com.tr

Özet: Bu çalışmada, ses komut tanıma sistemi ile gezgin araç kontrolü gerçekleştirilmiştir. Bu sistemde ses komut tanıma ortak vektör yaklaşımı (OVY) ile yapılmaktadır. OVY bir altuzay sınıflama yöntemidir. Bu yöntem ile her sınıfa ait değişmez özellikleri taşıyan ortak vektör elde edilir. OVY, yalıtık ses tanıma uygulamasında saklı Markov model (HMM) yöntemi ile karşılaştırılabilir sonuçlar vermektedir. Gezgin araç beş ses komutu ile kontrol edilmektedir: “ileri”, “geri”, “sağa dön”, “sola dön” ve “dur”. Deneysel çalışmada, OVY ile kişiye bağımlı ses komut tanımadaki yüksek başarımlar elde edilmiştir.

Anahtar Kelimeler: Ses Komut Tanıma, Ortak Vektör Yaklaşımı, Gezgin Araç Kontrolü.

Mobil Vehicle Control With Voice Command Recognition

Abstract: In this study, mobile vehicle control is implemented with voice command recognition system. In this system, common vector approach (CVA) is used to recognize voice commands. The CVA is a subspace classifier, and in the CVA method, a common vector that represents invariant features of class is obtained. CVA in isolated speech recognition gives comparative results with hidden Markov Models (HMM). The mobile vehicle is controlled with five voice commands: “forward”, “backward”, “turn right”, “turn left” and “stop”. In the experimental study, we obtained high success rates with CVA for speaker-dependent voice command recognition.

Keywords: Voice Command Recognition, Common Vector Approach, Mobil Vehicle Control.

1. Giriş

Ses tanıma sistemleri günlük hayatımızda birçok alanda kullanılmaya başlanmıştır. Ev, otomobil, ofis, hastane ve fabrika gibi çalışma ortamlarında ses tanıma ve sesle kontrol uygulamaları görmektediriz[1,8,15,16]. Ses komut sistemleri kullanıcıların insanlarla etkileşiminde, araç kontrolünde ve kişisel ve kamuya açık veri erişiminde kullanılmaktadır[16]

Ses tanıma; yalıtık, birleşik ve sürekli olmak üzere üç şekilde yapılmaktadır. Ses tanıma sistemleri, kişiye bağımlı veya kişiye bağımsız ses tanıma yapacak şekilde tasarlanabilir[4].

Ses komut tanıma sistemlerinde sesin bütünü

için bir model oluşturulabilir. Bu modelleme şekli özellikle sınırlı sayıda ses tanıma sistemleri için uygundur. Tanınacak ses sayısı fazla olduğu durumlar için ses modellemesi için ses birimlerinden yararlanılır. Ses birimleri ile ses/komut tanımadaki sürekli ses tanıma yöntemi uygulanır.

Ses komut tanıma sistemlerinde yalıtık veya sürekli konuşma tanıma yapılabilir [2,5,6,13,14]. Bu sistemlerde komutların değişken olması ve bu komutların sisteme kolay bütünleşmesi isteniyorsa sürekli konuşma tanıma yöntemleri tercih edilmelidir.

Günümüzde ses tanıma sistemlerinde “state-of-art” yöntemi olan saklı Markov modeller

(HMM) kullanılmaktadır[4-6,14]. HMM yanında deęişik ses tanıma yöntemleri de kullanılabilir. Bu çalışmada ses komut tanıma için bir altuzay sınıflama yöntemi olan OVY yöntemi kullanılmıştır[2,5,13]. OVY, başarımlar olarak HMM'e yakın sonuçlar vermektedir. Ayrıca OVY'nin HMM'e göre eğitim ve tanımda işlem açısından üstünlükleri vardır[7].

Bu çalışmada, OVY kullanılarak ses komut tanıma sistemi ile gezgin araç kontrolü yapılmıştır. Gezgin araç kontrolü için "ileri", "geri", "sağa dön", "sola dön" ve "dur" komutları kullanılmaktadır.

2. Ortak Vektör Yaklaşımı

OVY, ses ve görüntü tanıma uygulamalarında kullanılan bir altuzay sınıflama yöntemidir[3,9,11,12]. Bu yöntem ile her sınıfa ait deęişmez özellikleri taşıyan ortak vektör elde edilir. OVY hem yeterli veri durumu ($m \geq n$), hem de yetersiz veri durumları ($m < n$) için uygulanabilir [9,11,12]. Burada m ses komut sınıfına ait vektör sayısını, n ise her vektördeki eleman sayısını göstermektedir. Bu çalışmada OVY yetersiz veri durumu için ses komut tanıma yapılmaktadır.

OVY' de eğitim ile bir sınıfına ait ortak vektör ve o sınıfın farksızlık altuzayını veren özvektörler elde edilir. Yetersiz veri durumunda ($m < n$) ortak vektör ve farksızlık altuzayı Gram-Schmidt dikleştirme yöntemi veya altuzay teknikleri kullanılarak elde edilebilir. Altuzay tekniğinde ortak vektör ile ortak deęişinti matrisinin özvektörleri arasındaki ilişkiden yararlanılır[10].

Eğitim setinde bir sınıfına ait öznelik

vektörleri $a_1, a_2, \dots, a_m \in R^n$ ile gösterilsin

Bu sınıfa ait ortak deęişinti matrisi O Eşitlik (1)'den elde edilir.

$$\Phi = \sum_{i=1}^m (a_i - a_{ort})(a_i - a_{ort})^T \quad (1)$$

Eşitlik (1)'de aort, sınıf ortalama vektörünü

göstermektedir. OVY' de öznelik uzayı, farklılık altuzayı B ve farksızlık altuzayı B--olmak üzere birbirine dik iki altuzaya ayrılır. Farksızlık altuzayı B--, ortak deęişinti matrisinin sıfır özdeğerlerine karşılık gelen özvektörler tarafından gerilir. Yetersiz veri durumu ($m < n$) için ortak deęişinti matrisi O, $n-m+1$ adet sıfır özdeğere sahip olacaktır. Bu özdeğerler büyükten küçüğe sıralandığında fark ve farksızlık altuzayları için özdeğerler aşağıdaki şekilde seçilir:

$$\underbrace{\lambda_n \geq \lambda_{n-1} \geq \dots \geq \lambda_{n-m+2}}_{\substack{m-1 \\ \lambda_i \neq 0 \\ \text{Fark Altuzayı}}} \geq \underbrace{\lambda_{n-m+1} \geq \dots \geq \lambda_2 \geq \lambda_1}_{\substack{n-m+1 \\ \lambda_i = 0 \\ \text{Farksızlık Altuzayı}}}$$

λ_i özdeğerine karşılık gelen özvektör u_i ile

gösterilsin. Bu durumda farksızlık altuzay izdüşüm matrisi aşağıdaki şekilde bulunur.

$$P^\perp = \sum_{j=1}^{n-m+1} u_j u_j^T \quad (2)$$

Bir sınıfa ait ortak vektör a_{ortak} , o sınıfa ait

her hangi bir öznelik vektörünün farksızlık altuzayına izdüşümünden elde edilir:

$$a_{ortak} = P^\perp a_i \quad \forall i = 1, 2, \dots, m \quad (3)$$

OVY' de tanıma yapılırken en küçük öklid uzaklık ölçütü kullanılır. Bu ölçüte göre, bilinmeyen sese ait öznelik vektörü a_x 'in her sınıfın farksızlık altuzayındaki izdüşümünün, o sınıfın ortak vektörüne olan uzaklığı bulunur. Bilinmeyen ses, en küçük uzaklığı veren sınıfa atanır.

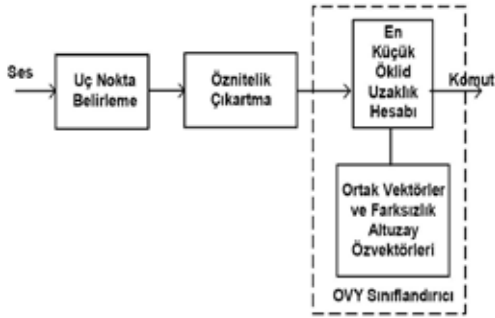
$$c^* = \arg \min_{1 \leq c \leq K} \|P^{c\perp} a_x - a^{c\perp}\| \quad (4)$$

Eşitlik (4)'de K, sınıf sayısını göstermektedir.

3. Ses Komut Tanıma Sistemi

Genel bir ses komut tanıma sisteminin blok şeması Şekil 1'de verilmektedir. Bu sistemde, bilinmeyen komuta ait sesin uç noktaları belirlendikten sonra öznitelik vektörü elde edilir. OVY ile öznitelik vektörünün hangi komuta ait olduğu belirlenir. Daha sonra bu komut sayısal I/O kartı ve RF verici üzerinden gezgin araca gönderilir.

Ses komut sisteminde tanıma yapılmadan önce ses komutlarının uç noktalarının belirlenmesi gerekir. Bu çalışmada uç nokta belirme için enerji ve sıfır geçiş ölçüleri kullanılmıştır [17]. Enerji ve sıfır geçiş sayısı için elde edilen eşik değerleri kullanılarak sesin uç noktaları belirlenmektedir.



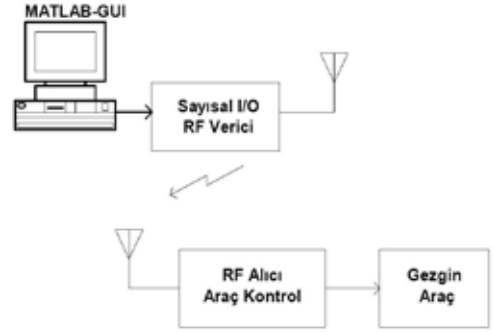
Şekil 1. Ses komut tanıma sistemi

Öznitelik çıkartma yapılırken önce ses örneklerine önvurgulama yapılmaktadır. Daha sonra ses örnekleri 32 ms'lik (256 örnek) çerçevelere ayrılarak çerçeveler üzerine Hamming pencereleme uygulanır. Çerçeveler arasındaki örtüşme 8 milisaniyedir (64 örnek). Her çerçeve için 12 MFCC (Mel Frequency Cepstral Coefficients) hesaplanmaktadır. Bu katsayıların ilk değeri logaritmik çerçeve enerjisini (LogE) vermektedir. Her çerçeve için elde edilen MFCC katsayıları arka arkaya eklenerek ses komutu için öznitelik vektörü ax elde edilmektedir.

4. Gezgin Araç Kontrolü

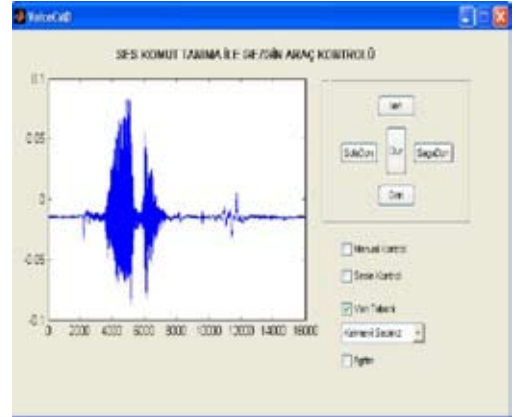
Ses komut tanıma ile gezgin araç kontrolünde kişisel bilgisayar, I/O kartı, RF alıcı ve verici

ve gezgin araç kullanılmaktadır (Şekil 2).



Şekil 2. Ses komut tanıma ile gezgin araç kontrol sistemi

Kişisel bilgisayar üzerinde MATLAB ortamında koşturulan yazılım ile ses komutları tanınmaktadır. Komut tanıma ile ilgili hareket bilgileri önce I/O kartına, oradan da RF verici ile gezgin araca iletilmektedir. Gezgin araç üzerindeki RF alıcı ile hareket bilgisi çözümlenmekte ve araç üzerindeki motora gerekli uyartım sinyalleri uygulanmaktadır.



Şekil 3. Sesli komut tanıma ve gezgin araç kontrolü için MATLAB- GUI.

Sesli komut tanıma ve gezgin araç kontrolü için MATLAB ortamında GUI hazırlanmıştır (Şekil 3). Bu GUI kullanılarak ses komutlarının kişiye bağımlı eğitimi yapılabilmektedir. GUI'de ses komutlarının modellenmesi için her komut için 10 kayıt yapılması istenmektedir. GUI ile her

kayıt grafiksel olarak çizilmekte ve bu kayıt sonra geri çalınmaktadır. Kullanılıcı bu kaydı kabul etmekte veya reddederek tekrar yeni bir ses kaydedebilmektedir.

Sistemdeki RF alıcı/vericinin ve gezgin aracın doğru çalıştığından emin olmak için GUI'ye elle kumanda seçeneği eklenmiştir. Bu seçenek ile GUI üzerinde komutlara karşılık gelen düğmeler üzerinden istenilen komut gezgin araca gönderilebilmektedir.

Bilgisayardan RF vericiye komutlar sayısal I/O kartı üzerinden gönderilmektedir. Bu çalışmada Advantech PCL818 veri toplama kartı kullanılmıştır.

GUI üzerinde ses kontrol seçeneği seçilerek gezgin aracın sesle kontrolü yapılmaktadır.

5. Sonuç

Bu çalışmada ses komut tanıma sistemi OVY ile gerçekleştirilmiş ve gezgin araç kontrolü yapılmıştır. Sistemin başarımı kişiye bağımlı tanıma için beklenildiği gibi %100 olarak elde edilmiştir.

MATLAB ortamında hazırlanan GUI, ses komut tanıma sisteminin geliştirilmesi için uygun bir ortam sağlamaktadır. Bu GUI ile farklı ses tanıma yöntemlerinin seçilmesi sağlanabilir. Bu şekilde farklı yöntemlerin ses komut tanıma uygulamasındaki başarımları karşılaştırılabilir.

Bu sistem, daha küçük ölçekli ve Digital Signal Processor (DSP) tabanlı ses komut tanıma sistemlerinin geliştirilmesinde kullanılabilir.

6. Kaynaklar

[1]. Abordo A., Liao J., "Voice Command Recognition: Robocard", Senior Project, University of California, Riverside Department of Electrical Engineering, June 2003.

[2]. CMU Sphinx Speech Recognizer, <http://sourceforge.net/projects/cmuspinx>.

[3]. Çevikalp, H., Neamatu, M., Wilkes, M., and Barkana, A., "Discriminative Common Vectors for Face Recognition", IEEE Transaction on Pattern Analysis and Machine Intelligence, 27: 4-13, 2005.

[4]. Deler J.R., Proakis J.G. and Hansen J.H., Discrete-Time Processing of Speech Signals, Macmillian, 1993. [5]. Dikte Konuşma Tanıma, <http://www.dikte.com.tr/> [6]. Dragon Naturally Speaking Software, <http://www.nuance.com/naturallyspeaking/>

[7]. Edizkan R., "Gizli Markov Model ile Bilgisayarda Konuşma Tanıma: Özellik Uzayında ve Altuzayda Sınıflandırıcı Tasarımı", Doktora Tezi, Eskişehir Osmangazi Üniversitesi Elektrik-Elektronik Mühendisliği Bölümü, 1999.

[8]. Gellatly A.W, The Use Of Speech Recognition Technology In Automotive Applications, Doktora Tezi, Virginia Polytechnic Institute and State University 1997.

[9]. Gülmezoğlu, M.B., Dzhafarov, V., Keskin, M. and Barkana, A., "A Novel Approach to Isolated Word Recognition", IEEE Trans. Speech and Audio Proc., 7(6):620-628, 1999.

[10]. Gülmezoğlu, M.B., Dzhafarov, V., and Barkana, A., "The Common Vector Approach and its Relation to Principle Component Analysis", IEEE Trans. Speech and Audio Proc., 9(6):655-662, 2001.

[11]. Gülmezoğlu, M.B., Dzhafarov, V., and Barkana A., "Örüntü Sınıflandırıcılar İçin Ortak Vektör Yaklaşımı", 7. Sinyal İşleme ve Uygulamaları Kurultayı, Haziran 1999.

[12]. Gülmezoğlu, M.B., Dzhafarov, V., Edizkan R. and Barkana A., "The common vector approach and its comparison with other subspace methods in the case of sufficient data", Elsevier Computer and Speech Language dergisinde Nisan 2007 basılacak.

- [13]. Günal, S., Edizkan, R., Barkana, A., “The design of real-time digit recognizer using the common vector approach”. In: Proc. Eleventh National Conf. Signal Processing and Applications. İstanbul, 2003, 308-311.
- [14]. GVZ, <http://www.gvz.com.tr/>.
- [15]. Hale, C.; Nguyen, C , “Voice command recognition using fuzzy logic”, 7-9 Nov. 1995, 608-613.
- [16]. Larson J.A. , Speech-enabled Appliances, <http://www.larson-tech.com/> November/December 2000.
- [17]. Rabiner, L.R., Sambur, M.R.,. “An algorithm for determining the endpoints for isolated utterances”. The Bell System Technical Journal, 1975, vol:54, 297-315.